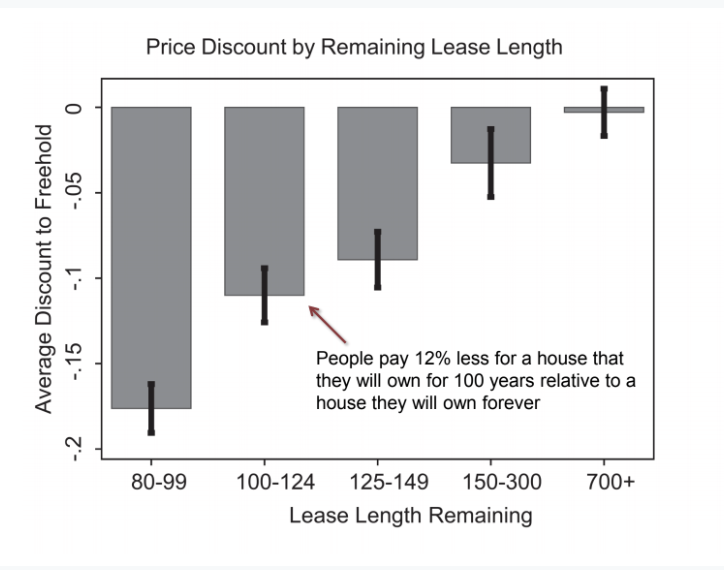
Lecture 15: Policies to Mitigate Climate Change

Professor Raj Chetty, Harvard University

# The Social Cost of Carbon

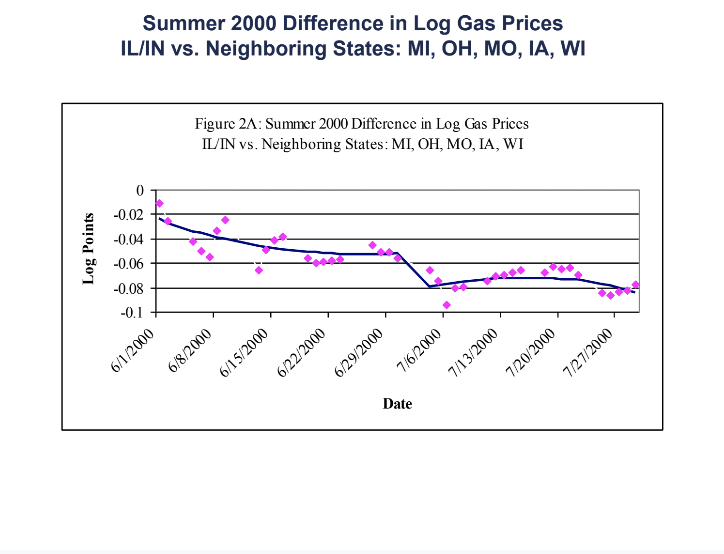
* When discussing policies to mitigate climate change, we first need to think about estimating the social cost of carbon. There are three steps in that process.
  + The first is to predict the impact of carbon emissions on the climate using a climate forecasting model. That is part of climate science and outside the scope of this class.
  + Second, we measure the impacts of the changes in climate on various things that affect human welfare (e.g. economic productivity, health consequences). In Lecture 14, for example, we looked at the impacts of the Clean Air Act on children’s health outcomes and future earnings.
  + Third, we want to measure the current social cost of climate change by figuring out how we should add and discount all future damages in terms of reduced economic output or worse health.
* The third step is non-trivial. Suppose we do not care at all about future generations. We only care about the present generation. Then climate change might have very little cost because if we think most of the impacts of climate change are going to come down the road, in 2100 or beyond, and we put very little weight on the future it is not going to be very costly.
* At the other extreme, if we take the view that we care equally about all generations, the current generation and all future generations that are going to be alive, then the costs are going to become infinite. If we have this 20% reduction in GDP year after year for an infinite number of years going forward, then the cost is infinitely large. The question is basically where do we fall between those two extremes? We have to have some way to weight a sequence of future payoffs.
* We care about discounted rates not just for next year relative to the present year, but we care about discount rates in the very long run over periods of hundreds of years. How can we try to figure that out using real transactions that occur in the world? A recent paper by Giglio, Maggiori and Stroebel develops an approach to do this using data on residential property sales in the UK and in Singapore in the 2000s.
  + There are two different types of contracts in these two countries.
  + There are what are called freeholds which are similar to the way we would buy houses in the United States which is perpetual ownership.
  + There is also a leasehold where someone can have ownership for a fixed period of time, often a very long period of time. They can buy the rights to a piece of land for 100 years or for 1,000 years. There are properties that basically get transferred back to the state after a certain point in time. As a result, owners have a finite period of ownership. What the paper does is compare how much people pay for two identical properties, one of which is a leasehold and one of which is a freehold, and thus implicitly calculate how buyers discount cash flows in the very long run.



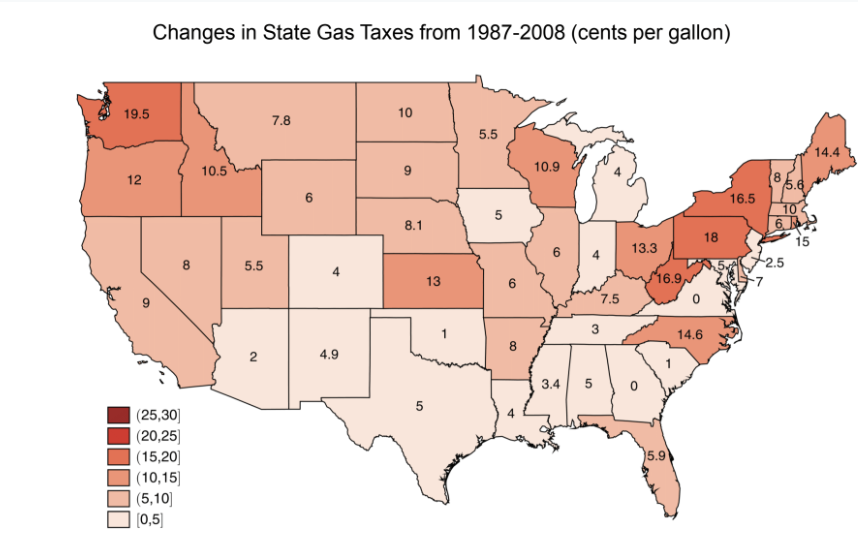
* The authors find that people will pay about 12% less for a house that they will own for 100 years relative to a house that they will own forever. This allows us to figure out how people value future cash flows.
  + In particular, the fact that people place a discount on houses that they can keep for only 100 years relative to houses that they can keep forever means implicitly that they place substantial value on money that they will have access to even 100 years from now.
  + That shows that the implied discount rate is not high enough to ignore cash flows 100 years from now. We can calculate an implied annual discount rate by saying if we value money 100 years from now at 12.6% less than the freehold, that implies given the stream of flows that we have an annual discount rate of 2.6%. That is to say a $1,000 a year from now would on average be worth about $974 today.
* Putting together all this literature that has developed over the years, a working group during the Obama administration estimated that the final social cost of carbon is $40 per ton of carbon dioxide that is emitted.
  + This is a very specific exact number that the government now uses in numerous policy decisions from fuel economy rules for cars to regulations on power plants.
  + However, while $40 number seems like a done-and-dusted answer, it is certainly not set in stone and remains highly debated both from a scientific point of view and a political point of view.
  + The Trump administration for instance has advocated using a 7% discount rate, instead of a 3% discount rate when calculating the social cost of carbon. That of course makes a huge difference because 7% when we have compounding basically means that cash flows 100 years from now are worth a lot less than cash flows at present. That drives the social cost of carbon down to $5 per ton instead of $40 per ton.
* At $5 per ton, the types of policies that make sense in terms of regulating vehicle emissions and other things look very different than at $40 per ton. Pinning down things like the discount rate is extremely important. Ultimately there are going to be some political and policy decisions that factor into this.

# How to Mitigate Climate Change

* The most common policy tool to try to change human behavior is corrective taxes, also known as Pigouvian taxes, that increase the private costs of consuming a good like gasoline in order to then ultimately reduce something like carbon emissions. Are these policies effective? What are their impacts? How should taxes be set in order to get carbon emissions to a level given the $40 per ton environmental cost?
* As is intuitive, taxes on gas are one obvious way to reduce gas consumption and ultimately carbon dioxide emissions. After levying a gas tax, one of two things can happen.
  + One possibility is that they are passed on to consumers. Consumers face higher prices at the pump and maybe they end up buying less gasoline.
  + A different possibility is that Exxon decides to keep the price of gas the same, and they end up just eating the higher tax in terms of having lower profits, in which case the prices that consumers pay at the pump might not actually change and the total amount of gasoline consumption might not change.
  + Figuring out the extent to which any given tax falls on consumers versus producers is a useful first step in understanding the potential behavioral impacts it might have. The answer is not obvious. If demand is highly inelastic, consumers are not very responsive to prices. More of the tax change is going to get passed on to the consumers.
* Doyle and Sampatharank who study this question in a paper using state-level gas tax reforms and a difference-in-differences design. In this particular case, gas prices spiked above $2 in 2000. In response to that, in order to make life easier for consumers, a number of states started to suspend their gas taxes.
  + In particular, Indiana suspended its gas tax on July 1st of 2008 and then reinstated it on October 30th. Similarly, Illinois suspended its gas tax on July 1st and reinstated it on December 31st. This creates nice variation from an econometric perspective because we can use Indiana as a treatment group and then compare it to nearby Midwestern states like Ohio which did not have this policy change. Illinois and Indiana suspended their gas tax.



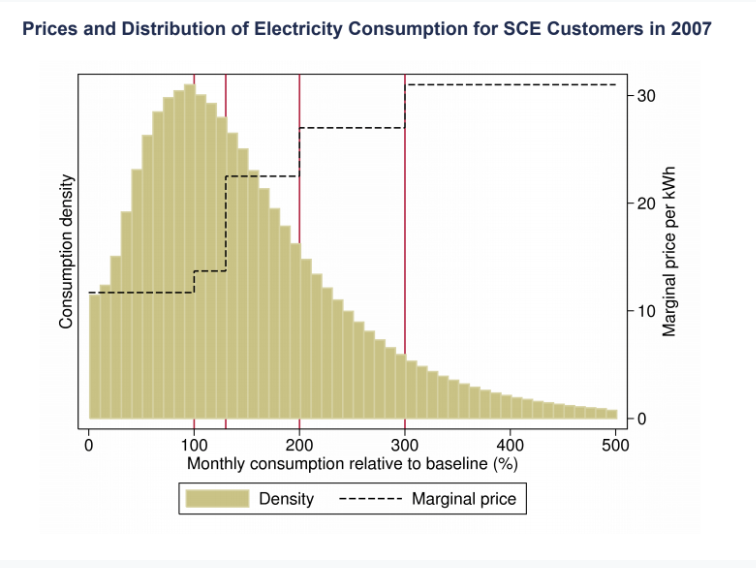
* The bottom line finding is that a 10-cent increase in the gas tax leads to a seven-cent increase in the price paid by consumers. About 70% of the incidence of gas taxes at the state level falls on consumers and 30% falls on producers. 30 cents of each dollar of the gas tax is coming out of the profits of oil firms and 70 cents is coming from consumers' pockets.
* If people are paying higher prices at the pump means that when we raise gas taxes, they could potentially reduce consumption of gasoline significantly. The next question is when we see this seven-cent increase in prices, how much less gas do people actually consume? Are people very inelastic in the sense that does not matter what the price of gas is, consumers drive the same amount?
* A paper by Li et al uses data covering all 50 states. They exploit changes in tax rates in all of those states over a much longer time period from 1966 to 2008. There is quite a bit of differential variation across states in gas taxes over time.



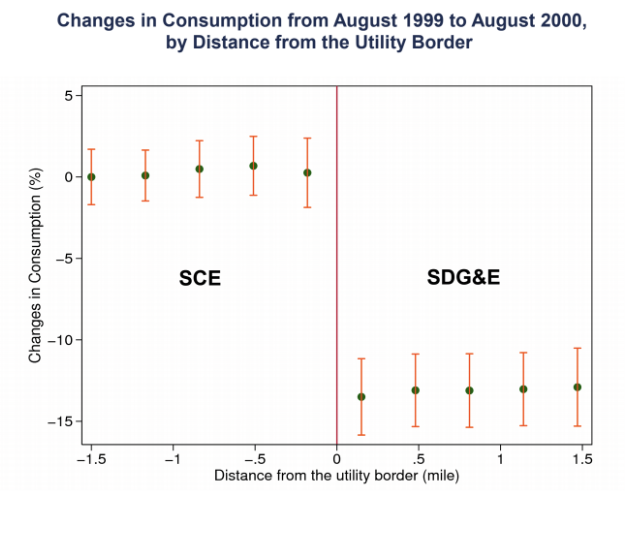
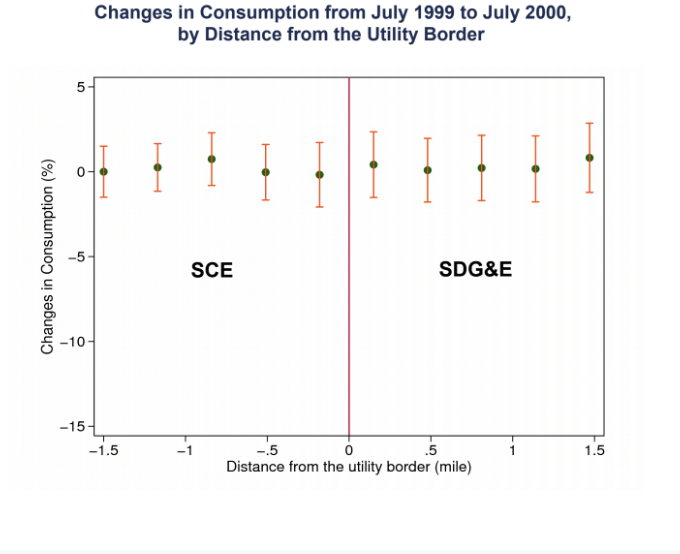
* That is the critical thing in order to be able to implement a difference-in-difference design. With just federal increases in gas taxes the design would fail because there are lots of other things that affect demand for gas over time like changes in public transit or preferences and so forth. Because we have this staggered variation across states, we can use that to identify the causal effects of changes in gas taxes on gasoline consumption basically by using each state as a control group for other states.
  + This generalization of the difference-in-difference approach to 50 states in 44 years yields more than 500 experiments of the type that we just saw with Illinois and Indiana using a method called fixed effects regression.
  + The key identification assumption is that gas consumption would have changed in the same way in any given state relative to the nation as a whole had there been no differential tax change.
* The results of the paper show that a 10% increase in the gas tax or 10% increase in the price due to a gas tax change will reduce gasoline consumption by about 4%. The elasticity is 0.4.
  + In other words, a 10-cent increase in the gas tax would reduce gasoline consumption by 1.7%. Gas taxes do matter. If we were to add a $1 tax to gasoline, in the US we would reduce gas consumption by about 17% which is quite a bit. Another thing to keep in mind is that this estimate is identified off of short-run variation—year to year. It is saying that in a given year of gas taxes rise by $1 people would consume 17% less gas, but the long run impact could be even bigger because if we have sustained high gas prices for the next 10 years some people might think about not buying a car as opposed to just driving the car they already have less. People may live closer to work or use public transit.
* In terms of ultimate carbon emissions, researchers have estimated that the transportation sector accounts for about one third of carbon emissions in the US.
  + That implies that a 10-cent increase in the gas tax which would reduce gas consumption by 1.7% and would reduce overall carbon emissions by one third of 1.7% which is about half a percent. Equally, a $1 gas tax would reduce carbon emissions by about 5% in the US. Scientists predict that we need to cut carbon dioxide emissions by about 50% if we want to stop the increase in global temperatures. 5% is non-trivial but nowhere near enough to completely stop global warming.

# Electricity Markets

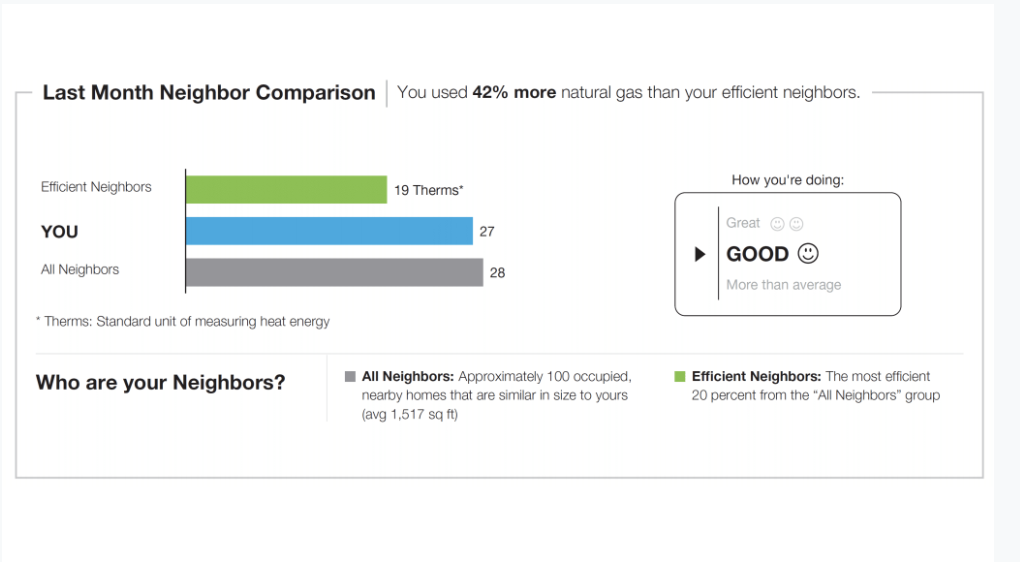
* Electricity is priced differently than gasoline. Electricity is often priced in tiers. The price of an additional kilowatt of electricity is higher when consumers have already consumed a lot of electricity.
  + If someone consumes a certain number of kilowatt hours, there is a certain rate they get charged. If they can consume even more than that, they get bumped into a higher tier and then an even higher tier and so forth.
* How responsive are people to higher electricity prices as they make choices about electricity consumption?
  + In principle we can examine the distribution of the outcome variable—the distribution of electricity consumption in this case—in order to understand the impacts of the tiered price schedule. In particular, at points where prices change when someone gets bumped into a next tier, we should expect to see bunching in the distribution of people who are responding.
* A paper by Koichiro Ito in 2014 studies the impacts of prices on electricity usage using household level billing data from utility companies in Orange County.
  + Specifically, he exploits the fact that the utility company that provides service depends upon where families live. The two electricity companies are Southern California Edison and San Diego Gas & Electric. SDG&E serves the southern side of the county and SCE serves the northern side.
  + A simple model would predict that if consumers are optimizing then they will turn off their AC with some probability if they are close to the tier-increasing threshold because it is not worth it to continue running the AC at that higher price level relative to the previous price level. We should see a bunch of spikes in the distribution at each of those cut-offs if people are responding to these tiered prices. The histogram that shows the actual distribution of electricity consumption is completely smooth. There is no evidence that people are responding to these tiered prices.



* There are two interpretations.
  + One possibility is that people are not actually aware that the price is changing. If we think about running the AC in a house, would the residents actually know when they crossed into the 150% tier versus not? That is a lack of salience issue.
  + An alternative explanation is that maybe consumer demand for electricity is actually insensitive to the price. Perhaps people actually do not care about the price that much and they want to run their AC or their fridge regardless and price does not matter.
  + These two very different interpretations have very different implications for policy. If people do not care about the price, then apparently increasing electricity prices is not going to impact energy usage and may not be a great way to try to curb energy usage. If it is about salience, then we might think about a very different set of interventions where we try to provide more information by making prices clearer to consumers.
* To distinguish between these two explanations, Ito uses a second strategy which is where we return to the fact that there are two electricity providers here.
  + He uses a spatial regression discontinuity approach. In summer of 2000, SDG&E raised its average electricity prices just across the board, they just raised the prices in a way that people would have heard about in the news whereas SCE did not change anything in the summer of 2000.
  + He uses a regression discontinuity design to estimate the effect of this change. To see how this works, imagine a set of households that are different distances from the border between the two firms’ operating zones.



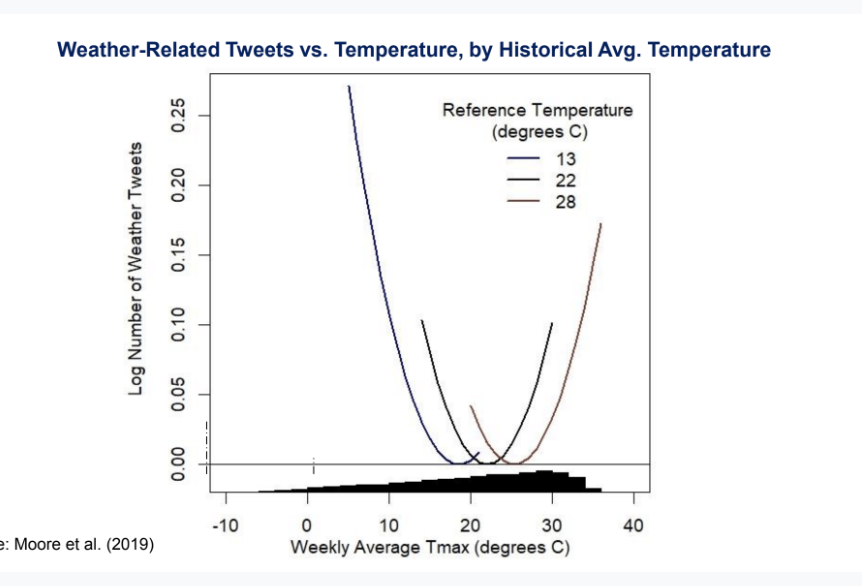
* + Minus 0.5 is a household that is a half mile on the side of the SCE. Plus 0.5 is a household that is a half mile into the SDG&E of the border. Over a one-year period before this policy change occurred, the change in electricity consumption was zero for households that were on the SCE side and the SDG&E side. Now doing the same analysis looking from August 1999 to August 2000, since SDG&E changed prices just before August of 2000 there is extremely clear evidence that people reduced electricity consumption in the southern side of Orange County by about 15% when prices went up relative to the people on the other side of the line where prices were held the same.
* That is fairly clear evidence pointing to the fact that people are responsive to electricity prices. However, with tiered prices, people are not aware when they are going into the next bracket.
  + Even though there is a system to try to discourage consumption, it does not work because people do not know about it. That implies that when we think about designing corrective taxes of the types that we are talking about here, salience and the structure of the incentives matters as much as the dollars involved.
* Traditionally when people think about policy design, they think about prices and not about how those prices are presented. Tiered pricing or changes in average prices should have the same effect in the traditional economic model where everybody's optimizing. However, it is very clear that that is not the case.
* There are two approaches using the learnings from above.
  + The most obvious thing to do is try to make prices more salient to consumers using things like smart meters. A company called Opower has been doing innovative technological work where owners can get meters in their houses that show which bracket they are in and the marginal cost of turning on the AC.
  + A different approach is once we admit that people are not perfectly optimizing and are influenced by things like information and other things beyond the price of the good, we might think about other non-price tools motivated by results from the social psychology literature for instance. Well-known work by Cialdini and collaborators demonstrates that social comparisons and injunctive social norms can have significant impacts on behavior in many contexts and in particular in the context of electricity use. In a typical bill using the results of the social norms research, residents can see how much they consumed, how much all of their neighbors consumed on average and how much their “efficient neighbors”—those in the bottom 20% of consumption—consumed. This last part is what social psychologists call an injunctive norm which is putting some value judgment on one’s consumption. There have been a series of randomized experiments evaluating whether these things actually work and the answer is that they do.
* The first intervention is sending out a flyer that only has the descriptive information, leaving out the smiley face and the good, bad, and great performance. The people who were above average in terms of electricity usage to begin with cut back. No one wants to be inefficient compared to their neighbors. Unfortunately, the people who are below average seem to increase the amount they consume because they feel entitled to turn on their AC more because they are consuming less than their neighbors. That is not great in terms of overall impacts.



* By adding a “grade” of good, great, bad, two things happen. First, the amount of reduction for the people who are above average becomes even larger. The second thing that seems to happen is that for below average consumers do not increase the amount they are consuming. The social norms treatment reduces electricity usage by about one kilowatt hour per day which is about a 2.5% reduction in electricity usage.

# Salience of Climate Change

* The last set of issues is to understand why—given the importance of climate change as an issue—public concern about climate change, and hence political support, might be limited.
  + There is a recent paper by Fran Moore and co-authors presents some evidence that is illuminating on this issue and also illustrates some nice methodology that we can use with modern data.
  + Their hypothesis is that because climate change is very gradual, temperatures go up slowly, climate is evolving steadily but in a gradual way over time, the remarkability or the salience of changes in climate gets attenuated because people perceive something that is very exceptional relative to what they have seen in the past. If something is changing very gradually, people are just less likely to notice it and as a result, less likely to react to it. In a sense gradual increases in temperature go unnoticed even though the total absolute change in the climate might be quite large. That is the hypothesis for why we might not see much support for policies to address changes in the climate.
* This paper presents evidence for that idea by looking at the adjustment of perceptions using data from Twitter to measure how much people discuss the weather.
  + They use 2.2 billion tweets between 2014 and 2016 which are geocoded. They have data on local area temperature measurements from 1981 to 2016. Area by area across the United States, they can see exactly how temperatures have been changing. Their goal is to estimate a model that relates the frequency of weather-related tweets to temperature anomalies, in particular testing whether when the temperature is very different from the recent past, if there is very different tweeting behavior.
* In order to implement this the first thing to do is to see if a set of tweets are weather related or not. Obviously with 2.2 billion tweets, we cannot do that by hand.
  + The authors use a fairly simple and standard approach in the machine learning literature which is called a bag of words classification algorithm. They define a tweet as weather-related if it contains any word in a list of predefined weather-related terms that we have identified from some source or based on own knowledge. They take a set of words that they define as having something to do with the weather. Then they search each of the 2.2 billion tweets and see if one of these words appears. If it appears in a tweet, label the tweet weather related, and if it does not appear, label that tweet not weather related. Naturally, some errors will occur.
  + To validate the approach, they take a sample of 6000 tweets where they manually define them as whether related or not. This is the “truth.” They compare this to the bag-of-words classification algorithm. Many things that people tweet about are related to the weather.
    - “It is too hot to be dressin' cute.”
    - “I swear when it comes to driving in the rain, people in Southern California are idiots on the road.”
* These are cases where the bag-of-words algorithm is working correctly. All of these tweets seem to be about the weather. Quite often, we get misclassified tweets.
  + “I need an ice cold Bud Light.”
* By the bag-of-words algorithm that would be a tweet classified as being about the weather. Obviously, it is not about the weather. There are a number of tweets like this which are false positives. They use the validation sample to figure out which ones are false positives and which ones are true positives.
  + In particular they calculate the rate of false positives and true positives. 46% of the tweets they classify as weather related are actually false positives.
  + That might seem like a bad thing, but a critical lesson in thinking about machine learning applications here and elsewhere is that it actually does not matter if we have errors and even if we have a significant number of errors. What matters is whether those errors are correlated with the hypothesis. In particular what they show here is that the errors are uncorrelated with temperature fluctuations. They do not see an excess number of the Bud Light type (false positives) tweets when the temperature is anomalously hot or cold. From an experimental perspective, the random noise is acceptable as long as it is balanced between the cases where the weather is unusually cold versus the temperature being unusually high.
* The authors use these tweets to get back to the original hypothesis and test whether perceptions of climate adjust slowly to changes in the temperature. The authors find that when the temperature becomes very low relative to a place where the average temperature is 13 degrees Celsius, there are a lot more weather-related tweets. There are about 20% more weather related tweets when the temperature is five degrees below the typical average. When exceptionally cold or exceptionally hot, people are talking more about the weather and the same thing holds true in places that are relatively hot.



* Interestingly, what actually affects tweeting frequency is not the absolute temperature but the deviation in the temperature from the typical norm in the area.
  + If it is 15 degrees, people are tweeting a lot more about the weather in a place where the average temperature is typically 22 degrees, not in a place where the average temperature is typically 13 degrees. This of course makes sense intuitively. It is more salient when the temperature deviates from what an average resident is typically used to. If There is a mild snowstorm in Boston, people are much less excited about that than if that happens in Atlanta.
* Next, the authors estimate a dynamic adjustment model where they relate tweets in a given week to the rate of temperature anomalies both in that same week and in prior weeks, in prior years. They are trying to see whether or not individuals get insensitive as it gets hotter and hotter. Does a given anomaly have less of an effect if there was a sequence of unusually hot temperatures in the past?
  + There is about a 10% increase in weather related tweets when the temperature is 3 degrees cooler, according to the paper. Suppose there is a three-degree colder week not just this year but also in the previous year and in the year before and then the year before that. The result is this decaying pattern such that if there is a sequence of five years in a row where it has been a little bit colder than average, people are no longer tweeting about it the way they were if it was just one year where the temperature is exceptionally cold. That shows that people's perceptions are adapting if it starts to become colder and colder year after year.